

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2000-174817

(43)Date of publication of application : 23.06.2000

(51)Int.Cl. H04L 12/56  
H04Q 3/00  
H04Q 11/04

(21)Application number : 11-172584 (71)Applicant : NEC CORP

(22)Date of filing : 18.06.1999 (72)Inventor : RAMAMURT  
GOPALAKR  
FAN RUIXU  
SMILJANIC  
ALEKSANDR

(30)Priority

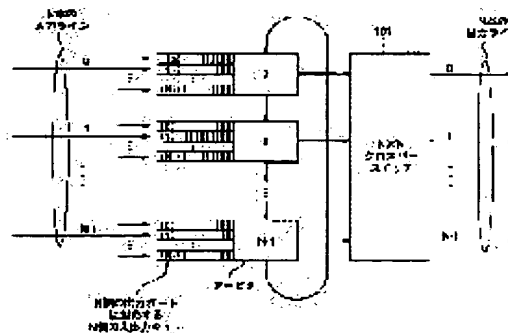
Priority number : 98 206975 Priority date : 08.12.1998 Priority country : US

## (54) METHOD AND SYSTEM FOR SCHEDULING IN EXCHANGE SYSTEM

(57)Abstract:

PROBLEM TO BE SOLVED: To allow a title system to satisfy a tight timing requirement by selecting one output from among a set of outputs that can be available for a future time slot, corresponding to a selected input and storing the selected input and the selected output relating to the selected input in pairs as a schedule.

SOLUTION: An ultrahigh speed exchange system consists of an  $N \times N$  crossbar switch 101, and packets received from  $N$ -sets of input lines consist of cells with a fixed length entirely. Each of the  $N$ -sets of input ports has  $N$ -sets of arbiters respectively



and each input port has N-sets of theoretical queues corresponding respectively to N-sets of output ports. Each of the N-sets of the arbiters (0, 1,..., N-1) receives a set of outputs of the available output ports from a preceding arbiter sequentially each time slot and selects one output port from among the available output ports according to the round robin system. The output of the selected output port is excluded from the set of outputs.

---

**LEGAL STATUS**

[Date of request for examination]	14.03.2000
[Date of sending the examiner's decision of rejection]	13.03.2001
[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]	
[Date of final disposal for application]	
[Patent number]	3221436
[Date of registration]	17.08.2001
[Number of appeal against examiner's decision of rejection]	2001-05660
[Date of requesting appeal against examiner's decision of rejection]	12.04.2001
[Date of extinction of right]	

Copyright (C); 1998,2003 Japan Patent Office

(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-174817

(P2000-174817A)

(43) 公開日 平成12年6月23日 (2000.6.23)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	ターミナル* (参考)
H 0 4 L 12/56		H 0 4 L 11/20	1 0 2 Z
H 0 4 Q 3/00		H 0 4 Q 3/00	
11/04		11/04	R

審査請求 有 請求項の数18 O L (全 16 頁)

(21) 出願番号	特願平11-172584	(71) 出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22) 出願日	平成11年6月18日 (1999.6.18)	(72) 発明者	ゴバラクリシナン ラママーシー アメリカ合衆国, ニュージャージー, 08540 プリンストン, インディペンデンス ウェイ 4, エヌ・イー・シー・ユ ー・エス・エー・インク内
(31) 優先権主張番号	09/206975	(74) 代理人	100097157 弁理士 桂木 雄二
(32) 優先日	平成10年12月8日 (1998.12.8)		
(33) 優先権主張国	米国 (US)		

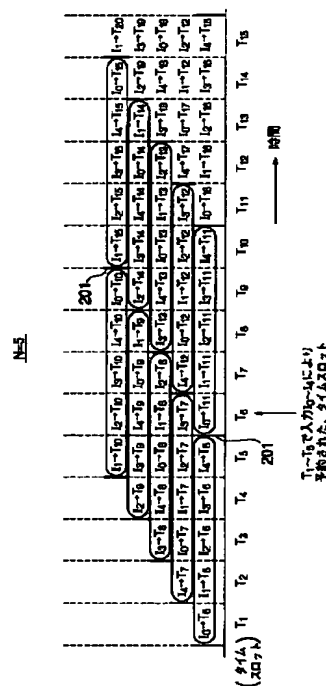
最終頁に続く

(54) 【発明の名称】 交換システムにおけるスケジューリング方法及び装置

(57) 【要約】

【課題】 超高速交換システムにおいて入出力間スケジュールを高速に確立することができる新規なスケジューリング方法及び装置を提供する。

【解決手段】  $N \times N$  交換システムにおいて、 $N$  入力にそれぞれ対応した論理的キューと、 $N$  入力にそれぞれ対応し隣接アービタへ制御情報を送出する  $N$  個のアービタとが設けられている。複数のスケジューリング過程がタイミングをずらせて開始され、 $N$  タイムスロットだけ未来のタイムスロットがそれぞれ決定される。開始アービタからラウンドロビン方式で順次1つのアービタが選択され、選択されたアービタは未来のタイムスロットにおける利用可能な出力の集合の中から1つの出力を選択する。選択された出力を出力集合から除外し、択された入力とそれに関連する選択された出力との組をスケジュールとしてメモリに記憶する。複数のスケジューリング過程をパイプライン処理により同時に並列処理できるために、スケジュールを高速に確立することができる。



(19) 日本国特許庁 (J P)

## (12) 公開特許公報 (A)

(11) 特許出願公開番号

特開2000-174817

(P2000-174817A)

(43) 公開日 平成12年6月23日 (2000.6.23)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テマコード* (参考)
H 0 4 L 12/56		H 0 4 L 11/20	1 0 2 Z
H 0 4 Q 3/00		H 0 4 Q 3/00	
11/04		11/04	R

審査請求 有 請求項の数18 O L (全 16 頁)

(21) 出願番号	特願平11-172584	(71) 出願人	000004237 日本電気株式会社 東京都港区芝五丁目7番1号
(22) 出願日	平成11年6月18日 (1999.6.18)	(72) 発明者	ゴバラクリシナン ラママーシー アメリカ合衆国, ニュージャージー, 08540 プリンストン, インディペンデン ス ウエイ 4, エヌ・イー・シー・ユ ー・エス・エー・インク内
(31) 優先権主張番号	09/206975	(74) 代理人	100097157 弁理士 桂木 雄二
(32) 優先日	平成10年12月8日 (1998.12.8)		
(33) 優先権主張国	米国 (US)		

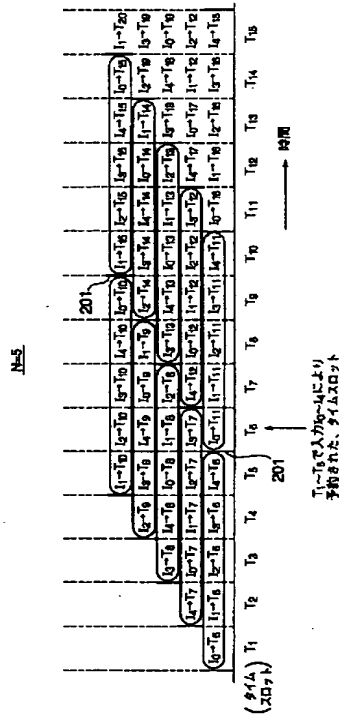
最終頁に続く

(54) 【発明の名称】 交換システムにおけるスケジューリング方法及び装置

## (57) 【要約】

【 과제】 초고속 교환 시스템에 있어 입출력간 스케줄을 고속에 확립한 것이 가능한 신규 스케줄링 방법 및 장치를 제공한다.

【 해결 수단】  $N \times N$  교환 시스템에 있어,  $N$  입력의 각각에  $N$  출력에 각각 대응한 논리적이다 큐와,  $N$  입력에 각각 대응하고 인접 아비타에 제어 정보를 송출한  $N$  개의 아비타가 마련되어 있다. 여러의 스케줄링 과정이 타이밍을 비켜 놓을 수 있고 시작되고,  $N$  타임 슬롯만 미래의 타임 슬롯이 각각 결정된다. 시작 아비타로부터 라운드 올새 방식으로 순차적으로 1 개의 아비타가 선택되고, 선택된 아비타는 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택한다. 선택된 출력을 출력 집합으로부터 제외하고, 택 된 입력과 그것에 관련된 선택된 출력과의 쌍을 스케줄로서 메모리에 기억한다. 여러의 스케줄링 과정을 파이프라인 처리에 의하고 동시에 병렬 처리할 수 있기 위해(때문에), 스케줄을 고속에 확립한 것이 가능하다.



#### 【특허 청구의 범위】

【청구항 1】 N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖고, 각 입력은 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적 이다 큐) 으로 된 교환 시스템에 있어서 스케줄링 방법에 있어,

- 임의의 스케줄링 과정을 시작한 입력에 대하여, 해당 시작 입력으로부터 미리 정해진 수의 타임 슬롯만 미래의 타임 슬롯을 결정하고,
- 상기 시작 입력으로부터 라운드 울새 방식으로 순차적으로 1 개의 입력을 선택하고,
- 선택된 입력에 전송해야 할 패킷이 존재하면, 해당 선택된 입력에 대하여, 상기 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택하고,
- 선택된 출력을 상기 출력 집합으로부터 제외하고, 상기 선택된 입력과 그것에 관련된 상기 선택된 출력과의 쌍을 스케줄로서 기억하다

스텝으로 된 것을 특징으로 한 스케줄링 방법.

【청구항 2】 여러의 스케줄링 과정이 파이프라인 처리에 의하고 동시에 진행한 것을 특징으로 한 청구항 1 기재된 스케줄링 방법.

【청구항 3】 N이 홀수의 경우, 상기 미래의 타임 슬롯은 상기 시작 입력으로부터 N 타임 슬롯만 미래에 위치하고, 상기 선택된 출력을 상기 출력 집합으로부터 제외한 출력 집합을 다음에 선택된 것이 당연한 입력에 전송하다, 것을 특징으로 한 청구항 1 기재된 스케줄링 방법

【청구항 4】 N 개의 스케줄링 과정이 1 타임 슬롯씩 위상을 비켜 놓으면서 동시에 진행한 것을 특징으로 한 청구항 3 기재된 스케줄링 방법.

【청구항 5】 N이 짝수의 경우, 상기 미래의 타임 슬롯은 상기 시작 입력으로부터 (N + 1) 타임 슬롯만 미래에 위치하고, 상기 선택된 출력을 상기 출력 집합으로부터 제외한 출력 집합을 다음에 선택된 것이 당연한 입력에 전송한 동작을 (N + 1) 타임 슬롯중 1 개 1 타임 슬롯분 지연시키다, 것을 특징으로 한 청구항 1 기재된 스케줄링 방법.

【청구항 6】 (N + 1) 개의 스케줄링 과정이 1 타임 슬롯씩 위상을 비켜 놓으면서 동시에 진행한 것을 특징으로 한 청구항 5 기재된 스케줄링 방법.

【청구항 7】 상기 교환 시스템은, 더욱, 멀티 캐스트 패킷을 격납하기 위한 별개의 큐를 갖고, 각 큐는 해당 큐의 HOL 패킷의 수신인을 나타내는 멀티 캐스트 비트맵(BM) 을 갖고,

멀티 캐스트 큐는 유니캐스트 큐보다도 우선되고 처리되고, 임의의 입력은 상기 출력 집합과 멀티 캐스트 비트맵과의 적집합을 충족시키는 모든 출력을 선택하고,

HOL 멀티 캐스트 패킷은, 상기 미래의 타임 슬롯에 있어 선택된 모든 출력에 송신되다,

것을 특징으로 한 청구항 1 내지 6 의 어느 한쪽에 기재된 스케줄링 방법.

【청구항 8】 N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖고, 각 입력은 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적 이다 큐) 으로 된 교환 시스템에 있어,

상기 N 입력의 각각은,

도착 패킷을 격납하고 상기 논리적이다 큐를 관리한 입력 격납 수단과,  
 상기 도착 패킷의 출력 요구에 따르고, 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택하고 상기 입력 격납 수단에 송출한 아비타 수단과,  
 상기 아비타 수단에 의하고 선택된 출력과 그 입력과의 쌍을 접속 정보로서 격납한 접속 격납 수단과,  
 상기 아비타 수단을 인접한 아비타 수단이라고 협조하고 제어한 파이프라인 제어 수단과,  
 로 되고,

상기 파이프라인 제어 수단은,

- 임의의 스케줄링 과정을 시작한 입력의 아비타 수단에 대하여, 해당 시작 아비타 수단으로부터 미리 정해진 수의 타임 슬롯만 미래의 타임 슬롯을 결정하고,
- 상기 시작 아비타 수단으로부터 라운드 올새 방식으로 순차적으로 1 개의 아비타 수단을 선택하고,
- 선택된 아비타 수단에 대응한 입력 격납 수단으로부터 출력 요구가 있으면, 해당 선택된 아비타 수단에 의하고, 상기 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택시키고,
- 선택된 출력을 상기 출력 집합으로부터 제외하고, 상기 선택된 아비타 수단과 그것에 관련된 상기 선택된 출력과의 쌍을 스케줄로서 상기 접속 격납 수단에 격납하다,  
 것을 특징으로 한 스케줄링 장치.

【청구항 9】 N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖는  $N \times N$  크로스 버스 근질거림을 갖는 교환 시스템에 있어,  
 상기 N 입력의 각각은,

도착 패킷을 격납하고 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적이다 큐) 관리한 입력 모듈과,  
 상기 도착 패킷의 출력 요구에 따르고, 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택하고 상기 입력 격납 수단에 송출한 아비타 와,  
 상기 아비타에 의하고 선택된 출력과 그 입력과의 쌍을 격납한 접속 메모리와,  
 상기 아비타를 인접한 아비타라고 협조하고 제어한 파이프라인 제어부와,  
 로 되고,

상기 N 출력의 각각은 출력 모듈로 되고,

상기 파이프라인 제어 수단은,

- 여러의 스케줄링 과정을 순차적으로 시작한 아비타에 대하여, 해당 시작 아비타로부터 미리 정해진 수의 타임 슬롯만 미래의 타임 슬롯을 각각 결정하고,
- 상기 시작 아비타로부터 라운드 올새 방식으로 순차적으로 1 개의 아비타를 선택하고,
- 선택된 아비타에 대응한 입력 모듈로부터 출력 요구가 있으면, 해당 선택된 아비타에 의하고, 상기 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택시키고,
- 선택된 출력을 상기 출력 집합으로부터 제외하고, 상기 선택된 아비타와 그것에 관련된 상기 선택된 출력과의 쌍을 스케줄로서 상기 접속 메모리에 격납하고,

상기 스케줄에 따라 입력 모듈의 패킷을 상기 크로스바 스위치를 통하여 선택된 출력에 송출한 것을 특징으로 한 교환 시스템.

【청구항 10】 N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖고, 각 입력은 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적이다 큐) 으로 된 라운드-올새-글리 티-스케줄링 프로토콜을 위한 크로스바 스위치에 있어서 타임 슬롯 결정 방법이고, 상기 프로토콜의 입력은 모든 입출력 큐의 상태이고, 상기 프로토콜의 출력은 스케줄이고,

- $i = (\text{정수} - k - 1) \bmod N$ 에 대응한 입력을 선택하고,
- 만약 입력이 없으면 정지하고, 그 밖이라면 라운드 올새의 방식으로  $i = (i + 1) \bmod N$ 에 의하고 결정된 다음 입력을 선택하고,
- 집합  $C = \{ (i, j) \mid \text{출력 } j \text{에 대응한 입력 } i \text{에 있어 적어도 1 개의 패킷이 존재하다} \}$ 의 요소인 쌍  $(i, j)$ 가 존재한다면, 출력 j를 선택하고,
- 스텝 c)에 있어 상기조  $(i, j)$ 가 존재하지 않는다면, 입력 집합으로부터 i를 제거하고 스텝 b)에 돌아오고,
- 입력 집합으로부터 i를, 출력 집합으로부터 j를 각각 제거하고,
- 상기조  $(i, j)$ 를 상기 스케줄에 가하고 스텝 b)에 돌아오다,

스텝으로 된 것을 특징으로 한 타임 슬롯의 결정 방법.

【청구항 11】 각 타임 슬롯에 있어, N 개가 다른 스케줄이 미래의 N 타임 슬롯 사이에서 동시 진행한 스케줄링 방법에 있어,

- 라운드 올새 방식으로, 스케줄링을 위한 입력에 대하여 특정한 미래의 타임 슬롯을 이용 가능하게 하여,
- 입력 i에 의하고, 미래의 k 번째의 타임 슬롯에 대한 출력을 선택하고,

c) 미래의  $k$  번째의 타임 슬롯에 대한 스케줄을 시작하고,

d) 다음 입력  $(i + 1) \bmod N$ 을 결정하고, 상기  $k$  번째의 타임 슬롯의 사이에 패킷을 자유롭게 수신할 수 있는 나머지 출력을 상기 다음의 입력에 송출한다,

시스템으로 된 것을 특징으로 한 스케줄링 방법.

【청구항 12】 입력  $i$ 가  $k$  번째의 타임 슬롯에 대한 스케줄을 완료하지 않았으면, 출력을 선택하고, 갱신된 출력 집합을 다음 입력에 송출하고, 상기 입력  $i$ 가 상기  $k$  번째의 타임 슬롯에 대한 스케줄을 완료한다면, 상기 출력 집합을 상기 다음의 입력에 송출하지 않다, 것을 특징으로 한 청구항 11 기재된 방법.

【청구항 13】 출력의 갱신 집합을 전의 입력으로부터 수신하지 않는 입력은, 새로운 스케줄을 시작한 것을 특징으로 한 청구항 12 기재된 방법.

【청구항 14】 입력이 홀수개의 경우의 파이프라인·라운드 올새·구리디·스케줄링 방법에 있어,

e)  $k(i, h) > 0$ 은 입력  $i$ 가  $h$  번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - h) \bmod N$ 은  $h$  번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타내고,  $k(i, h) = 0$ 은  $h$  번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한 것으로 하고,  $k(0,1) = k(1,1) = \dots = k(N-1,1) = 0$ , 정수  $= N+1$ 에 초기화하고,

f)  $O_{h+N} = \{0,1,\dots,N-1\}$ ,  $0 \leq h \leq N-1$  및  $i \neq i_h$ 로서,  $k(i_h, h) = h+N$  및  $k(i, h) = k((i-1) \bmod N, h-1)$ 라고(와) 설정하고,

g)  $0 \leq i \leq N-1$ 인 입력  $i$ 로 패킷을 송신해야 할 출력  $j$ 를 집합  $O_{k(i,h)}$ 로부터 라운드 올새 방식으로 선택하고 (고치고  $k(i,h) \neq 0$ 이다),  $j$ 를  $O_{k(i,h)}$ 로부터 제외하고,

h)  $0 \leq i \leq N-1$ 인 입력  $i$ 로 선택된 출력  $j$ 의 어드레스를 커넥션 메모리의 메모리 위치  $k(i,h) \bmod N$ 에 기억하고, 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 이동하고,

i)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$ 로, 다음 입력  $(i+1) \bmod N$ 에 집합  $O_{k(i,h)}$ 를 전송하고,

j)  $0 \leq i \leq N-1$ 인 상기 입력  $i$ 와, 입력  $i$ 의 메모리 로케이션  $(h \bmod (N+1))$ 로부터 판독된 어드레스의 출력과의 사이에 크로스바 커넥션을 확립하고,

k)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄링 된 송신 입출력 큐의 선두로 유지된 패킷을 스위치 코어를 통하여 송신한다, 시스템으로 된 프로세스를 이용하고  $h$  번째의 타임 슬롯을 완료시키는 방법.

【청구항 15】 입력이 짝수개의 경우의 파이프라인·라운드 올새·구리디·스케줄링 방법에 있어,

m)  $k(i, h) > 0$ 은 입력  $i$ 가  $h$  번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - h) \bmod N$ 은  $h$  번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타내고,  $k(i, h) = 0$ 은  $h$  번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한 것으로 하고,  $k(0,1) = k(1,1) = \dots = k(N-1,1) = 0$ , 정수  $= N+1$ 에 초기화하고,

n)  $O_{h+N+1} = \{0,1,\dots,N-1\}$ ,  $0 \leq h \leq N-1$ 인  $i$ 로 집합  $\{i_h, (i_h + 1) \bmod N\}$ 의 요소가 아님다면 하여,  $k(i_h, h) = h + N + 1$ ,  $k(\bmod(i_h + 1) \bmod N, h) = k(i_h, h-2)$ , 및  $k(i, h) = k((i-1) \bmod N, h-1)$ 에 설정하고,

o)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어 패킷을 송신해야 할 출력  $j$ 를 집합  $O_{k(i,h)}$ 로부터 라운드 올새 방식으로 선택하고 (단,  $k(i,h) \neq 0$ 이다),  $j$ 를  $O_{k(i,h)}$ 로부터 제외하고,

p)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어, 선택된 출력  $j$ 의 어드레스를 커넥션 메모리의 메모리 위치  $k(i,h) \bmod (N+1)$ 에 기억하고, 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 이동시키고,

q)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$ 에 있어,  $(i_h - 2) \bmod N$ 의 입력이 집합  $O_{k((i_h-2) \bmod N, h)}$  1 타임 슬롯 부분만 지연시켰던 후, 다음 입력  $(i+1) \bmod N$ 에 집합  $O_{k(i,h)}$ 를 전송하고,

r)  $0 \leq i \leq N-1$ 인 입력  $i$ 와, 입력  $i$ 의 메모리 위치  $(h \bmod (N+1))$ 로부터 판독된 어드레스의 출력과의 사이에 크로스바 커넥션을 확립하고

s)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄 된 송신 입출력 큐의 선두의 패킷을 스위치 코어를 통하여 송신한다, 시스템으로 된 프로세스를 이용하고  $h$  번째의 타임 슬롯을 완료시키는 방법.

【청구항 16】 멀티 캐스트·스케줄링이 라운드 올새·구리디·스케줄링·알고리즘에 편입되고 있고, 멀티 캐스트 패킷은 도착 순서로 처리된 방식으로 격납되고, 유니캐스트 큐보다도 우선적으로 처리되고,  $h$  번째의 타임 슬롯에 있어서 시스템은, 더욱,

u)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어,  $j \in O_{k(i,h)}$   $BM_{i,j}$  충족시키는 모든 출력  $j$ 를 선택하고, HOL 멀티 캐스트 패킷을  $k$  번째의 타임 슬롯으로 선택된 출력에 송신하고,

v)  $O_{k(i,h)} \cap BM_{i,j}$  공 집합이라면 유니캐스트 큐를 처리하고, 그 밖의 경우는, 선택된 출력을  $O_{k(i,h)}$  및  $BM_{i,j}$ 로부터 제외하고,

w)  $BM_{i,j}$  하nil이라면, HOL 멀티 캐스트 패킷을 멀티 캐스트 큐로부터 삭제하다,

시스템으로 된 것을 특징으로 한 청구항 14 기재된 방법.

【청구항 17】 멀티 캐스트 스케줄링이 라운드 올새·구리다·스케줄링·알고리즘에 편입되고 있고, 멀티 캐스트 패킷은 도착 순서로 처리된 방식으로 격납되고, 유니캐스트 큐보다도 우선적으로 처리되고,  $h$  번째의 타임 슬롯에 있어서 스텝은, 더욱,

x)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어,  $j \in O_{k(i,h)}$   $BM_{i,j}$  충족시키는 모든 출력  $j$ 를 선택하고, HOL 멀티 캐스트 패킷을  $k$  번째의 타임 슬롯으로 선택된 출력에 송신하고,  
 y)  $O_{k(i,h)} \cap BM_{i,j}$  공 집합이라면 유니캐스트 큐를 처리하고, 그 밖의 경우는, 선택된 출력을  $O_{k(i,h)}$  및  $BM_{i,j}$ 로부터 제외하고,  
 z)  $BM_{i,j}$  하위이라면, HOL 멀티 캐스트 패킷을 멀티 캐스트 큐로부터 삭제한다,  
 스텝으로 된 것을 특징으로 한 청구항 15 기재된 방법.

【청구항 18】 스테이지  $i$ 가 입력  $l$ 에 관련되고, 상기 스테이지  $i$ 는 미래의 타임 슬롯으로의 출력에의 송신을 스케줄링 하여, 상기 미래의 타임 슬롯은 모든 스테이지를 통하여 순차적으로 빗나가고 가는  $N \times N$  스위치를 스케줄링하기 위한  $N$  스테이지·파이프라인 시스템에 있어,

입력에 대응한 모든 파이프라인 스테이지는, 2개의 입력이 동시에 동일한 미래의 타임 슬롯을 선택하지 않도록, 동시에 스케줄링을 실행하고, 출력 슬롯은 라운드 올새 방식에 근거하고 선택되고, 출력이 있는 스테이지에 의하고 선택된 때, 해당 출력은, 파이프라인 스테이지가 입력에 의하고 있는 타임 슬롯으로 이미 선택된 출력을 선택하지 않도록, 출력의 프리 폴로부터 제거되다, 것을 특징으로 한  $N$  스테이지·파이프라인 시스템.

【발명의 자세한 내용한 설명】

【0001】

【발명이 속한 기술 분야】 본 발명은 전자 및 광학 미디어용의 어플리케이션으로 사용된 초고속 (테라비트 등급) 교환 시스템에 관계되고, 특히, 초고속 교환 시스템에 있어서 입출력간의 스케줄링 및 그 스케줄링을 실현하기 위한 교환 시스템에 관한다.

【0002】

【종래의 기술】 광대역화의 요구에 수반하고, 테라비트스이칭의 필요성이 점점 높아지고 오고 있고, 이와 같은 고속 스위칭에 관한 제안이 많이 이루어지고 있다. 베샤이 및 미인터 (M. Beshai and E. Miinter) 에 의한다 "Multi-tera-bit/s switchbased on burst transfer and independent shared buffers," ICC'95, pp.1724-1730, 마케오운 등 (N. McKeown et al) 에 의한다 "The tiny tera: a packet switch core," IE EE Micro, vol.171, Jan.-Feb. 1997, pp.26-33, 및 존, 시마주, 츠쿠다 및 유키마츠 (W. D. Zhong, Y. Shimazu, M.Tsukuda, and K. Yukimatsu) 에 의한다 "A modular Tbit/s TDM-WDM photonic ATM switch using optical buffers," IEICE Transactions on Communications, vol.E77-B,no.2, Feb. 1994, pp.190-196 을(를) 참조하여 주십시오.

【0003】 전자적으로 제어된 광스위치(switch) 스이칭코아 (견해를 바꾸면 논리적이다 크로스바 스위치) 는 고성능 스위치의 후보로서 매력적이다. 10Gbps의 라인 속도로, 64 바이트의 셀 / 패킷이 40ns 이내에 처리된 필요가 있기 때문이다.

【0004】 당업자에 있어 중요한 문제의 1개는, 이와 같은 광스위치(switch) 스이칭코아를 효율적으로 사용한 스케줄링의 고속 결정법의 확립이다. 이 경우, 스위치 (교환기) 의 설계에 있어, 입력 버퍼링, 출력 버퍼링 또는 그것들 양쪽을 이용한 것이 가능하다. 출력 버퍼링을 이용한 스위치로는, 출력 버퍼의 액세스 속도가 스위치 전체의 단위 시간당 처리량을 상회하고 있는 필요가 있다.

【0005】 이와 같은 출력 버퍼에 필요로 된 속도를 억제하기 위해(때문에), 녹아웃 법이 채용된다. 이 방법으로는, 유한개의 셀만이 출력 버퍼에 받아들여지고, 잔여물은 폐기된다. 빛 녹아웃 스위치는, 존, 시마주, 츠쿠다 및 유키마츠 (Zhong, Y. Shimazu, M.Tsukuda, and K. Yukimatsu) 에 의한다 "A modular Tbit/s TDM-WDM photonic ATM switch using optical buffers," IEICE Transactions on Communications, vol. E77-B, no.2, Feb. 1994, pp.190-196로 제안되고 있다. 빛 녹아웃 스위치로는, 각 출력이 몇 개인가의 광역Banyan 망과 빛 버퍼를 필요로 하기 위해(때문에) 구성이 복잡으로 된다.

【0006】 입력 버퍼를 이용한 스위치는 좀더 효율적으로 버퍼를 사용할 수 있고, 메모리의 대역폭도 라인 속도의 2 배로 끝난다. 간단한 방법으로는, 각 입력이 그 입력의 대기 행렬 (큐) 에 있어서 처음의 패킷을 송신한 요구를 낸다. 만약 2 이상의 입력이 동일한 출력 포트를 요구한 경우, 그 중의 1개가 랜덤하게 선택된다. 이 방법은, 캐롤 등 (M. J. Karol, M.G. Hiuchy, and S.P. Morgan) "Input vs. output queuing on a space-division packet switch," IEEE Transactions on Communications, vol. COM-35, no.12, Dec. 1987, pp. 1347-1356에 나타나고 있고, 이 입력 버퍼링 알고리즘에 의하고, 균일 교통 상황에서 0.587의 단위 시간당 처리량을 얻을 수 있고 있다. 교통이 균일지 않다는 경우에는, 단위 시간당 처리량은 더욱 저하된다.

【0007】 다른 몇 개인가의 스케줄링 법에 있어서는, HOL (Head Of Line : 라인의 선두) 패킷 이외의 패킷이 출력 포트를 요구한다. 팬, 아키야마 및 타나카(R. Fan, M. Akiyama, and Y. Tanaka)에 의한다 "An input buffer-typeATM switching using schedule comparison," Electronics and Communications in Japan: Part I, vol.74, no.11, 1991, pp.17-25; 모토야마, 페타 및 서리 (S. Motoyama, D.W. Petr, and V.S. Frost)에 의한다 "Input-queuedswitch based on a scheduling algorithm," Electronics Letters, vol.31, no.14, July 1995, pp.1127-1128; 오바라(H. Obara)에 의한다 "Optimum architecture for input queuing ATM switches," Electronics Letters, vol.27, no.7, March 1991, pp.555-557을 참조하여 주십시오.



【0008】보다(부터) 상세하게 말하면, 각 타임 슬롯에 있어, 1개의 입력은 여러의 출력 포트를 요구한다. 타임 슬롯마다 정확하게 4개의 요구를 내면, 1에 가까운 효율을 얻을 수 있다. 그렇지만, 이 방법으로는, 고속 시, 스케줄링을 위한 여러의 요구 및 그 확인 응답을 1 타임 슬롯 안에서 처리한 것이 가능하지 않는다(여기에서 1 슬롯은 패킷 송신시간을 나타낸다). 또, 교통에 급격한 변화가 있는 경우에는, 각 입력이 어느 출력을 요구하든지 을(를) 독립하고 결정한 것으로 스위치 퍼포먼스는 저하된 것이다.

【0009】스위치 퍼포먼스를 개선한 차면, 스위치 컨트롤러가 모든 입출력 버퍼의 큐를 알고 있으면 된다. 이 정보에 의하고, 스위치 컨트롤러는 각 타임 슬롯으로의 동시 송신수를 증가시키는 것이 가능하다. 그렇지만, S L I P 프로토콜에 있어서는, 여러의 출력이 독립하고 여러의 입력에 허가를 주기 때문에, 비효율적적이 되어 있다. 자세한 것은, N. McKeown et al. "Scheduling cells in an input-queued switch," Electronic Letters, vol.29, no.25, Dec. 1993, pp.2174-2175를 참조하여 주십시오.

【0010】또, 알고리즘에 의하고 입력간보다 좋은 협조를 달성하고 있는 예로서는, D. Guo, Y. Yemini, Z. Zhang, "Scalable high-speed protocols for WDM optical star networks," IEEE INFOCOM'94.을(를) 참조하여 주십시오. 단, 이 알고리즘은 스케줄링을 결정하는데도 많은 타임 슬롯을 필요로 한 점에서 불리하다.

【0011】또, 양호한 퍼포먼스를 나타내는 랜덤·구리다·스케줄링 법(RGS)이 제안되고 있다. 이것에 관해서는, R. Chipalkatti, Z. Xhang, and A. A. Acampora, "Protocols for optical star-coupler network using WDM: performance and complexity study," IEEE Journal on Selected Areas in Communications, vol.11, no.4, May 1993, pp.579-589, 및 D. Guo, Y. Yemini, Z. Zhang, "Scalable high-speed protocols for WDM optical star networks," IEEE INFOCOM'94를 참조하여 주십시오.

【0012】이 종래의 RGS 법으로는, 입력 및 그것에 대응한 매칭 한 출력의 양쪽이 랜덤하게 선택된다. 그렇지만, 실제로는, 이와 같은 랜덤 선택을 실장한 것은 곤란하다. 각 타임 슬롯에 있어, N 개의 패킷이 N 개의 입력으로부터 N 개의 출력에 전송되고 얻는 것에 주의되고 싶다.

#### 【0013】

【발명이 해결할 것 같다고 한 과제】상술했던 것처럼, 광스위치(switch) 스이칭코야를 효율적으로 사용하기 위한 스케줄링 결정 방법으로서 아직 만족해야 할 것이 없다. 즉, 입출력 버퍼링을 이용한 것으로는, 구성의 복잡화나 스위치의 단위 시간당 처리량의 저하를 일으킨다. 또, 스위치 컨트롤러가 입출력 버퍼의 큐를 집중 관리한 방식으로는, 스위치의 입력수 및 출력수가 증대한 에 수반하고 계산양이 현저하게 증대하고, 1 타임 슬롯 안에서 처리를 완료한 것이 곤란해진다.

【0014】그러면, 본 발명의 목적은, 초고속 교환 시스템에 있어 입출력간 스케줄을 고속에 확립한 것이 가능한 신규 스케줄링 방법 및 장치(라운드 올새·구리다·스케줄링 : 이하, RRGs라고 말한다.)를 제공한 것에 있다.

【0015】본 발명의 다른 목적은, 교환기의 내부 속도를 상승시키는 일 없게 엄격한 타이밍 요구를 충족시킴과 동시에 양호한 퍼포먼스를 얻을 수 있는 신규 파이프라인 아키텍처를 갖는 교환 시스템을 제공한 것에 있다.

#### 【0016】

【과제를 해결하기 위한 수단】본 발명에 의한 스케줄링 방법은, N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖고, 각 입력은 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적이다 큐) 으로 된 교환 시스템에 있어, a) 임의의 스케줄링 과정을 시작한 입력에 대하여, 해당 시작 입력으로부터 미리 정해진 수의 타임 슬롯만 미래의 타임 슬롯을 결정하고, b) 상기 시작 입력으로부터 라운드 올새 방식으로 순차적으로 1 개의 입력을 선택하고, c) 선택된 입력에 전송해야 할 패킷이 존재하면, 해당 선택된 입력 대하여, 상기 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택하고, d) 선택된 출력을 상기 출력 집합으로부터 제외하고, 상기 선택된 입력과 그것에 관련된 상기 선택된 출력과의 쌍을 스케줄로서 기억하다, 스텝으로 된 것을 특징으로 한다. 더욱, 여러의 스케줄링 과정이 파이프라인 처리에 의하고 동시에 진행한 것을 특징으로 한다.

【0017】N이 홀수의 경우에는, 상기 미래의 타임 슬롯은 상기 시작 입력으로부터 N 타임 슬롯만 미래에 위치하고, 상기 선택된 출력을 상기 출력 집합으로부터 제외한 출력 집합을 다음에 선택된 것이 당연한 입력에 전송한다.

【0018】N이 짝수의 경우에는, 상기 미래의 타임 슬롯은 상기 시작 입력으로부터 (N + 1) 타임 슬롯만 미래에 위치하고, 상기 선택된 출력을 상기 출력 집합으로부터 제외한 출력 집합을 다음에 선택된 것이 당연한 입력에 전송한 동작을 (N + 1) 타임 슬롯중 1 개 1 타임 슬롯분 지연시킨다.

【0019】본 발명에 의한 타임 슬롯 결정 방법은, N 입력 및 N 출력 (N은 2 이상의 정수) 을 갖고, 각 입력은 상기 N 출력에 각각 대응한 N 개가 논리적 대기 행렬 (논리적이다 큐) 로 된 라운드·올새·글리 디·스케줄링 프로토콜을 위한 크로스바 스위치에 있어서 타임 슬롯 결정 방법이고, 상기 프로토콜의 입력은 모든 입출력 큐의 상태이고, 상기 프로토콜의 출력은 스케줄이고,

a)  $i = (\text{정수} - k - 1) \bmod N$ 에 대응한 입력을 선택하고,

b) 만약 입력이 없으면 정지하고, 그 밖이라면 라운드 올새의 방식으로  $i = (i + 1) \bmod N$ 에 의하고 결정된 다음 입력을 선택하고,

c) 집합  $C = \{ (i, j) \mid \text{출력 } j \text{에 대응한 입력 } i \text{에 있어 적어도 1 개의 패킷이 존재하다} \}$ 의 요소인 쌍  $(i, j)$  이 존재한다면, 출력 j를 선택하고,

d) 스텝 c) 에 있어 상가조  $(i, j)$  가 존재하지 않는다면, 입력 집합으로부터 i를 제거하고 스텝 b) 에 돌아오고,

e) 입력 집합으로부터  $i$ 를, 출력 집합으로부터  $j$ 를 각각 제거하고,

f) 상기조 ( $i, j$ )를 상기 스케줄에 가하고 스텝 b)에 돌아오다,  
스텝으로 된 것을 특징으로 한다.

【0020】또, 본 발명에 의한 스케줄링 방법은, 각 타임 슬롯에 있어,  $N$ 개가 다른 스케줄이 미래의  $N$ 타임 슬롯 사이에서 동시 진행한 스케줄링 방법에 있어,

a) 라운드 올새 방식으로, 스케줄링을 위한 입력에 대하여 특정한 미래의 타임 슬롯을 이용 가능하게 하여,

b) 입력  $i$ 에 의하고, 미래의  $k$ 번째의 타임 슬롯에 대한 출력을 선택하고,

c) 미래의  $k$ 번째의 타임 슬롯에 대한 스케줄을 시작하고,

d) 다음 입력  $(i+1) \bmod N$ 을 결정하고, 상기  $k$ 번째의 타임 슬롯의 사이에 패킷을 자유롭게 수신할 수 있는 나머지 출력을 상기 다음의 입력에 송출하다,

스텝으로 된 것을 특징으로 한다.

【0021】본 발명에 의한 입력이 홀수개의 경우의 파이프라인·라운드 올새·구리디스케줄링 방법은,

e)  $k(i, h) > 0$ 은 입력  $i$ 가  $h$ 번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - h) \bmod N$ 은  $h$ 번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타내고,  $k(i, h) = 0$ 은  $h$ 번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한 것으로 하고,  $k(0,1) = k(1,1) = \dots = k(N-1,1) = 0$ , 정수  $= N+1$ 에 초기화하고,

f)  $O_{h+N} = \{0,1,\dots,N-1\}$ ,  $0 \leq i \leq N-1$  및  $i \neq i_h$ 로서,  $k(i_h, h) = h+N$  및  $k(i, h) = k((i-1) \bmod N, h-1)$ 라고(와) 설정하고,

g)  $0 \leq i \leq N-1$ 인 입력  $i$ 로 패킷을 송신해야 할 출력  $j$ 를 집합  $O_{k(i,h)}$ 로부터 라운드 올새 방식으로 선택하고 ( $k(i,h) \neq 0$ 이다),  $j$ 를  $O_{k(i,h)}$ 로부터 제외하고,

h)  $0 \leq i \leq N-1$ 인 입력  $i$ 로 선택된 출력  $j$ 의 어드레스를 커넥션 메모리의 메모리 위치  $k(i,h) \bmod N$ 에 기억하고, 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 이동하고,

i)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$ 로, 다음 입력  $(i+1) \bmod N$ 에 집합  $O_{k(i,h)}$ 를 전송하고,

j)  $0 \leq i \leq N-1$ 인 상기 입력  $i$ 와, 입력  $i$ 의 메모리 로케이션  $(h \bmod (N+1))$ 로부터 판독된 어드레스의 출력과의 사이에 크로스바 커넥션을 확립하고,

k)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄링 된 송신 입출력 큐의 선두로 유지된 패킷을 스위치 코어를 통하여 송신하다,  
스텝으로 된 프로세스를 이용하고  $h$ 번째의 타임 슬롯을 완료시키는 것을 특징으로 한다.

【0022】더욱, 본 발명에 의한 입력이 짝수개의 경우의 파이프라인·라운드 올새·구리디스케줄링 방법은,

m)  $k(i, h) > 0$ 은 입력  $i$ 가  $h$ 번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - h) \bmod N$ 은  $h$ 번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타내고,  $k(i, h) = 0$ 은  $h$ 번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한 것으로 하고,  $k(0,1) = k(1,1) = \dots = k(N-1,1) = 0$ , 정수  $= N+1$ 에 초기화하고,

n)  $O_{h+N+1} = \{0,1,\dots,N-1\}$ ,  $0 \leq i \leq N-1$ 인  $i$ 로 집합  $\{i_h, (i_h + 1) \bmod N\}$ 의 요소가 아님다면 하여,  $k(i_h, h) = h + N + 1$ ,  $k((i_h + 1) \bmod N, h) = k(i_h, h-2)$ , 및  $k(i, h) = k((i-1) \bmod N, h-1)$ 에 설정하고,

o)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어 패킷을 송신해야 할 출력  $j$ 를 집합  $O_{k(i,h)}$ 로부터 라운드 올새 방식으로 선택하고 (단,  $k(i,h) \neq 0$ 이다),  $j$ 를  $O_{k(i,h)}$ 로부터 제외하고,

p)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어, 선택된 출력  $j$ 의 어드레스를 커넥션 메모리의 메모리 위치  $k(i,h) \bmod (N+1)$ 에 기억하고, 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 이동시키고,

q)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$ 에 있어,  $(i_h - 2) \bmod N$ 의 입력이 집합  $O_{k((i_h-2) \bmod N, h)}$  1 타임 슬롯 부분만 지연시켰던 후, 다음 입력  $(i+1) \bmod N$ 에 집합  $O_{k(i,h)}$ 를 전송하고,

r)  $0 \leq i \leq N-1$ 인 입력  $i$ 와, 입력  $i$ 의 메모리 위치  $(h \bmod (N+1))$ 로부터 판독된 어드레스의 출력과의 사이에 크로스바 커넥션을 확립하고

s)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄 된 송신 입출력 큐의 선두의 패킷을 스위치 코어를 통하여 송신하다,  
스텝으로 된 프로세스를 이용하고  $h$ 번째의 타임 슬롯을 완료시키는 것을 특징으로 한다.

【0023】더욱, 멀티 캐스트 스케줄링이 라운드 올새·구리디스케줄링·알고리즘에 편입되고 있고, 멀티 캐스트 패킷은 도착 순서로 처리된 방식으로 격납되고, 유니캐스트 큐보다도 우선적으로 처리되고,  $h$ 번째의 타임 슬롯에 있어서 스텝은, 더욱,

u)  $0 \leq i \leq N-1$ 인 입력  $i$ 에 있어,  $j \in O_{k(i,h)} \cap \text{BM}_i$ 를 충족시키는 모든 출력  $j$ 를 선택하고, HOL 멀티 캐스트 패킷을  $k$ 번째의 타임 슬롯으로 선택된 출력에 송신하고,

v)  $O_{k(i,h)} \cap \text{BM}_i$ 가 공 집합이라면 유니캐스트 큐를 처리하고, 그 밖의 경우는, 선택된 출력을  $O_{k(i,h)}$  및  $\text{BM}_i$ 로부터 제외하고,

w) BM

1가 하늘이라면, HOL 멀티 캐스트 패킷을 멀티 캐스트 큐로부터 삭제한다, 스텝으로 된 것을 특징으로 한다.

【0024】본 발명에 의한 N 스테이지·파이프라인 시스템은, 스테이지 i 가 입력 l 에 관련되고, 상기 스테이지 i 는 미래의 타임 슬롯으로의 출력에의 송신을 스케줄링 하여, 상기 미래의 타임 슬롯은 모든 스테이지를 통하여 순차적으로 빗나가고 가는  $N \times N$  스위치를 스케줄링 하기 위한 N 스테이지·파이프라인 시스템이고, 입력에 대응한 모든 파이프라인 스테이지는, 2 개의 입력이 동시에 동일한 미래의 타임 슬롯을 선택하지 않도록, 동시에 스케줄링을 실행하고, 출력 슬롯은 라운드 올새 방식에 근거하고 선택되고, 출력이 있는 스테이지에 의하고 선택된 때, 해당 출력은, 파이프라인 스테이지가 입력에 의하고 있는 타임 슬롯으로 이미 선택된 출력을 선택하지 않도록, 출력의 프리 폴로부터 제거된 것을 특징으로 한다.

【0025】본 발명에 의하면, N 입력으로부터 라운드 올새 방식으로 순차적으로 1 개의 입력을 선택하고, 그 선택된 입력에 대하여 미래의 타임 슬롯에 있어서 이용 가능한 출력의 집합의 중(속)에서 1 개의 출력을 선택하고, 스케줄로서 기억한다. 이것에 의해, 여러의 출력 할당 동작(스케줄링)을 병렬 처리한 것이 가능해지고, 내부 속도의 향상을 필요로 하지 않고 엄격한 타이밍 기준을 충족시키는 것이 가능하다. 더욱, 병렬 처리에 의하고, RGS의 양호한 퍼포먼스를 손상시키지 않고, 100%에 가까운 사용율을 달성한 것이 가능하다.

【0026】

【발명의 실시의 형태】그림 1 은 본 발명에 의한 초고속 교환 시스템의 개념적 아키텍처를 나타내는 블록도이다. 본 교환 시스템은  $N \times N$  크로스바 스위치 10-1로 되고, N 책의 입력 라인으로부터 수신한 패킷은 전부 고정 길이의 셀이다. N 개의 입력 포트는 각각 N 개의 아비타를 갖고, 더욱 각 입력 포트는 N 개의 출력 포트에 각각 대응한 N 개의 논리 큐를 갖는다. 후술하는 바와 같이, N 개의 아비타 및 각 아비타에 대응한 N 개의 논리 큐는 RRG S 알고리즘에 의하고 제어된다. 이하, RRG S 스케줄링에 관하여 상세히 설명한다.

【0027】라운드 올새·구리다·스케줄링 (RRGS)

RRGS 프로토콜의 입력은, 모든 입출력 큐의 상태이다. 크로스바 스위치 10-1의 각 입력 포트를 i ( $i \in \{0, 1, \dots, N-1\}$ )라고 기록하면, 이와 같은 입력, 즉 입출력 큐, 은 다음과 같은 집합 C에 의하고 나타내는 것이 가능하다.

【0028】 $C = \{(i, j) \mid \text{출력 } j \text{에 대응한 입력 } i \text{에 있어 적어도 1 개의 패킷이 존재한다}\}.$

【0029】RRGS 프로토콜의 출력은, N 개의 입력을 N 개의 출력에 대응시키는 스케줄이다. 이와 같은 스케줄의 집합 S는 다음과 같이 나타내는 것이 가능하다.

【0030】 $S = \{(i, j) \mid \text{입력 } i \text{로부터 출력 } j \text{에 패킷이 보내진다}\}.$

【0031】당업자에 있어서는 분명하지만, 각 타임 슬롯에 있어, 1 개의 입력은 1 개의 패킷만을 송신할 수 있고, 1 개의 출력은 1 개의 패킷만을 수신할 수 있다. 이 조건하에서, 임의의 k 번째의 타임 슬롯의 스케줄은, 다음 스텝 1) ~ 4)에 의하고 결정된다.

【0032】스텝 1) 전입력의 집합을  $I_k = \{0, 1, \dots, N-1\}$ , 전 출력의 집합을  $O_k = \{0, 1, \dots, N-1\}$ 라고 한다.  $i = (\text{정수} - k - 1) \bmod N$ 의 계산식에 의하고, 즉 (정수 - k - 1)을 N로 제산한 잉여로서 입력 i를 선택한다. 스케줄을 시작한 입력을 이와 같이 선택한 것으로 스케줄링의 인푸리멘테이션을 간단하게 할 수 있는 것이다.

【0033】스텝 2) 만약  $I_{k+1}$ 가 하늘이라면, 정지한다. 하늘이 아니면, 라운드 올새의 방식으로  $i = (i + 1) \bmod N$ 으로 된 다음 입력 i를 선택한다.

【0034】스텝 3) (i, j) Ck와 같이  $O_k$ 로부터 출력 j를 라운드 올새의 방식으로 선택한다. 만약 그러한 출력이 존재하지 않는다면,  $I_k$ 로부터 i를 제거하고, 스텝 2)에 돌아온다.

【0035】스텝 4)  $I_k$ 로부터 입력 i를,  $O_k$ 로부터 출력 j를 제거하고,  $S_k$ 에 (i, j)를 가하고, 스텝 2)에 돌아온다.

【0036】상기 프로토콜은, 분명히 상술한 종래의 RGS 법의 개량이다. 상술했던 것처럼, 종래의 RGS 법으로는, 입력 및 그것에 대응한 매칭 한 출력의 양쪽이 랜덤하게 선택되고 있다. 그러나, 실제로는, 이와 같은 랜덤 선택을 실행한 것은 어렵다. 각 타임 슬롯에 있어, N 개의 패킷이 N 개의 입력으로부터 N 개의 출력에 전송되고 얻는 것에 주의되고 싶다.

【0037】이것에 대하여, RRG S로는, 있는 주어진 타임 슬롯의 스케줄링 프로세스는 N 개의 위상으로 된다. 있는 타임 슬롯의 스케줄링의 각 위상에 있어, 1 개의 입력은 해당 타임 슬롯으로의 송신을 위해 사용 가능한 잔존한 출력의 1 개를 선택한다. 후술하는 바와 같이(그림 4 참조), 1 개의 위상은, 입력 모듈(IM)으로부터 라운드 올새(RR)아비타에 출선 리퀘스트(요구)를 송출하고, RR아비타에 있어 라운드 올새 방식으로 출력 선택을 행하고, 그리고 RR아비타로부터 입력 모듈 IM에 요구 응답(출선 선택 응답)을 송출한, 라고 말한 스텝으로 된다. 입력이 출력을 선택한 라운드 올새(RR)순서는, 모든 입력에 대하여 같은 액세스를 보증하도록 타임 슬롯마다 순환적으로 시프트한다.

【0038】홀수개 입력의 파이프라인 RRG S

10Gbps일 것인 고속 링크 속도에 있어서는, N 위상을 1 개의 타임 슬롯(64 바이트의 패킷 사이즈를 가정한다면 40ns) 안에서 종료시키는 것은 할 수 있지 않는다. 링크 속도가 고속화한 것에 따라, 종래의 기술으로는 1 타임 슬롯으로 1 위상을 초과한 처리를 완료한 것이 가능하지 않는다.

【0039】이 문제를 해결하기 위해(때문에), 본 발명으로는 파이프라인 방식을 이용한다. 즉, 타임 슬롯마다, 미래의  $N$  개의 타임 슬롯에 걸쳐,  $N$  개가 다른 스케줄이 동시 진행된다. 있는 1 개의 스케줄의 각 위상은 1 개의 입력만큼 관련된다. 바꾸어 말하면, 임의의 타임 슬롯에 있어, 다른 여러의 입력은, 다른 다른 미래의 타임 슬롯에 대한 스케줄의 위상을 실행한 것이 된다.

【0040】정의: 있는 미래의 타임 슬롯  $T_{k\text{위}}$  스케줄이 완료됐다고는,  $N$  위상 전부가 완료된 때, 즉, 타임 슬롯  $T_k$ 에 있어 송신하기 위한 출력을 선택한 찬스 (성공 / 실패에 관계되지 않고) 가 모든 입력에 부여받았던 때, 점고 유.

【0041】있는 스케줄의  $N$  개의 위상을 완료하는데도  $N$  개의 타임 슬롯이 필요하지만, 후술하는 바와 같이 파이프라인 어프로치를 이용하고  $N$  개의 스케줄을 병렬 계산한 것에 의하고,  $N$  개가 다른 스케줄의  $N$  개의 위상을 1 타임 슬롯 안에서 완료시키는 것이 가능하다. 그러나, 이것은 타임 슬롯마다 1 개의 스케줄을 완료시키는 것과 결과적으로 등가이다.

【0042】바꾸어 말하면, RRGs에 있어서는, 있는 미래의 타임 슬롯이 모든 입력에 대한 라운드 올새 스케줄링에 이용 가능해진다. 즉, 미래의  $k$  번째의 타임 슬롯  $T_{k\text{위}}$  스케줄을 시작한 입력  $i$  는 라운드 올새 방식으로 출력을 선택하고,  $(i+1) \bmod N$ 인 다음 입력  $i$ 에 대하고, 해당  $k$  번째의 타임 슬롯  $T_{k\text{위}}$  패킷을 자유롭게 수신 가능한 출력 포트를 나타내는 집합  $O_{k\text{위}}$  송출한다.

【0043】전의 입력  $(i-1) \bmod N$ 로부터  $k$  번째의 타임 슬롯  $T_{k\text{위}}$  아직 이용 가능한 출력 포트의 집합  $O_{k\text{위}}$  수취한 임의의 입력  $i$ 는, 만약 가능하면, 이 중에서 1 개의 출력을 선택하고, 만약  $k$  번째의 타임 슬롯  $T_{k\text{위}}$  스케줄이 완료되지 않는다면, 갱신된 출력 집합  $O_{k\text{위}}$  다음 입력  $i = (i+1) \bmod N$ 에 송출한다.  $k$  번째의 타임 슬롯  $T_{k\text{위}}$  스케줄링이 완료됐다면, 갱신된 집합  $O_{k\text{위}}$  다음 입력  $(i+1) \bmod N$ 에 송출되지 않는다. 이것에 의해, 현실 타임 슬롯에 있어 출력 집합  $O_{k\text{위}}$  수취하지 않았던 입력  $(i+1) \bmod N$ 은, 다음 타임 슬롯으로 새로운 미래의 타임 슬롯을 위한 새로운 스케줄링을 시작한 것이다.

【0044】RRGS의 상기 스텝 (1) 은,  $N$  개의 타임 슬롯의 기간에 1 도, 입력이 출력 집합  $O_{k\text{위}}$  송출하지 않는 것을 함의하고 있다. 집합  $O_{k\text{위}}$  송출하지 않는 입력  $i$ 는  $k$  번째의 타임 슬롯에 있어서 출력을 선택한 최후의 입력이다.

【0045】정리 1 입력 수  $N$ 이 홀수로,  $(\text{정수} - k) \bmod N$ 의 입력이  $k-1$  번째의 타임 슬롯에 있어 집합  $O_{k\text{위}}$  송출하지 않는다면,  $(\text{정수} - k) \bmod N$ 의 입력은  $k$  번째의 타임 슬롯의 스케줄을 완료시킨다.

【0046】(증명) 상기 정리 1 은 다음 것을 함의한다.

- ① 각 타임 슬롯에 있어,  $N$  개 모든 입력은, 미래가 있는 타임 슬롯으로의 송신을 스케줄링 한 기회를 갖는다.
- ② 각 타임 슬롯에 있어, 입력은 미래의 1 을 초과하지 않는 타임 슬롯으로의 송신을 스케줄링할 수 있다.
- ③ 각 타임 슬롯에 있어, 있는 출력은 1 개의 입력만로부터 송신을 점수하도록 스케줄 될 것 같다.

【0047】 $(\text{정수} - k) \bmod N$ 의 입력  $i$  를  $k-1$  번째의 타임 슬롯에 있어 집합  $O_{k\text{위}}$  송출하지 않는 입력이라고 가정하면, 그 이전의  $N-1$  개의 입력의 각각은  $k$  번째의 타임 슬롯의 예약을 한 때에 집합  $O_{k\text{위}}$  송출하지 않으면 안된다. 또한,  $(k-1-j)$  번째의 타임 슬롯에 있어,  $(i+j) \bmod N$ 의 입력은 집합  $O_{k\text{위}}$  다음 입력에 송출하지 않는다. 또,  $(i-j) \bmod N$ 의 입력은  $k$  번째의 타임 슬롯을 위한 예약을 행한다. 이와 같은 스케줄은,  $1 \leq j \leq (N-1)$  의(것) 임의의  $j$  에 관하여,

$$(1 \leq j \leq (N-1)) \quad i - j \neq i + j \bmod N$$

$$(1 \leq j \leq (N-1)) \quad 2j \neq 0 \bmod N$$

$N$ 이

홀수이다라면, 실현 가능하다.

【0048】따라서  $k$  번째의 타임 슬롯의 스케줄이  $k-N$  번째의 타임 슬롯으로의  $(i+1) \bmod N$ 의 입력에 의하고 시작됐다고 말한 것은,  $k-N-1$  번째의 타임 슬롯으로의 입력  $i$ 가 집합  $O_{k-N\text{위}}$  송출하지 않았던 것을 의미한다.

(증명 중)

【0049】그림 2는, 홀수 입력수의  $5 \times 5$  크로스바 스위치를 이용한 경우의 RRGs 스케줄링의 일례를 나타내는 타이밍 차트이다. 그림 2로는, 5 개의 입력  $I_0 \sim I_4$ 와, 그것들 입력이 출력 포트를 선택한 타임 슬롯  $T_1, T_2 \dots$ 라는 관계가 나타나고 있다.

【0050】그림 2에 있어, 예를 들면 타임 슬롯  $T_{5\text{위}}$ , 입력  $I_{1\text{위}}$  타임 슬롯  $T_{10\text{위}}$  송신을 행하기 위한 출력 포트의 선택 (스케줄링)을 행하고, 입력  $I_{3\text{위}}$  타임 슬롯  $T_9$ 에 있어서 스케줄링을 행하고 있다. 또, 다음 타임 슬롯  $T_{6\text{위}}$ 으로는, 입력  $I_{1\text{위}}$  타임 슬롯  $T_8$ 에 있어서 스케줄링을 행하고 있다. 이하 마찬가지이다. 그림중에 참조 번호 201로 예시된 붉은 종선은, 그 전의 입력으로 스케줄링이 완료되고, 다음 입력으로 새로운 스케줄링이 시작한 것을 나타내고 있다. 관련된 타임 슬롯에 있어 출력을 선택한 최후의 입력인 경우에는, 해당 입력은 다음 입력에 집합  $O$ 를 송출하지 않는다. 이 조건은  $N=5$  타임 슬롯마다 발생하기 때문에, 입력은 모듈로  $N$  카운터에 의하고 집합  $O$ 를 송출하지 않는 것을 결정한 것이 가능하다.

【0051】마지막으로,  $h$  번째의 타임 슬롯 (예를 들면 현재의 타임 슬롯)에 있어서 RRGs의 동작을 설명한다.  $O_{k\text{위}}$ 는  $k$  번째의 타임 슬롯의 이용 가능한 출력의 집합을 나타낸다.  $k(i, h) > 0$ 은, 입력  $i$ 가  $h$  번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - h) \bmod N$ 은  $h$  번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타낸다. 또,  $k(i, h) = 0$ 은,  $h$  번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한다. 스케줄러는 적절한 초기화가 필요하고, 초기화 기간은  $N$  타임 슬롯 계속된다. 초기화가 처음의 타임 슬롯  $T_{1\text{위}}$

시작된다고 가정하면, 초기화는  $k(0,1) = k(1,1) = \dots = k(N-1,1)=0$ , 정수= $N+1$ 이라고 설정한 것에 의하고 시작된다. 즉, 그 이후 갱신될 때까지 처음의  $N$  타임 슬롯은 모든 입력의 동작이 억제된다. 또, 패킷은, 입력 포트에 있어, 논리적으로 분리한 큐 (입출력 큐) 로 대가한다고 가정한다. 이 입출력 큐로는 각 출력 포트에 대응해 1 개의 큐가 설치되기 위해(때문에), HOL (Head Of Line : 라인 선두) 블로킹이 방지된다. 더욱, 수신 입출력 큐 및 송신 입출력 큐도 마련되어 있다.

【0052】 그림 1 에 나타내는 시스템에 입각해서 설명하면, 다음과 같게 정식화할 수 있다.

【0053】 1)  $O_{h+N} = \{0,1,\dots,N-1\}$ ,  $0 \leq i \leq N-1$  및  $i \neq i_h$ 로서,  $k(i_h, h) = h + N$  및  $k(i, h) = k((i-1) \bmod N, h-1)$ .

【0054】 2)  $0 \leq i \leq N-1$ 인 입력  $i$ 는, 패킷을 송신해야 할 출력  $j$ 를 출력 집합  $O_{k(i,h)}$ 로부터 RR 방식으로 선택한다. 단,  $k(i,h) \neq 0$ 이다. 또한,  $0 \leq i \leq N-1$  및  $i \neq i_h$ 인 입력  $i$ 는 선행한 타임 슬롯에 있어  $(i-1) \bmod N$ 의 입력으로부터 집합  $O_{k(i,h)}$ 를 수신하고 있다. 선택된 출력  $j$ 는, 수신한 집합  $O_{k(i,h)}$ 로부터 제외된다.

【0055】 그림 1 에 나타내는 시스템에 있어,  $N$  개의 아비타 ( $0, 1, \dots, N-1$ )의 각각은, 타임 슬롯마다 순차적으로전의 아비타로부터, 이용 가능한 출력 포트의 정보, 즉 집합  $O_{k(i,h)}$ 를 수취하고, 그 이용 가능한 출력 포트의 중(속)에서 1 개의 출력 포트를 RR 방식으로 선택한다. 선택된 출력 포트는 집합  $O_{k(i,h)}$ 로부터 제외된다. 이렇게 갱신된 집합  $O_{k(i,h)}$ 는, 후술하는 바와 같이 다음 아비타에 전송된다.

【0056】 3)  $0 \leq i \leq N-1$ 인 입력  $i$ 는, 선택된 출력  $j$ 의 어드레스를 커넥션 메모리의 메모리 위치  $k(i,h) \bmod N$ 에 기억한다. 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 옮겨진다.

【0057】 4)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$  (현실 아비타)는, 다음 입력  $(i+1) \bmod N$  (후속한 아비타)에 집합  $O_{k(i,h)}$ 를 전송한다. 여기에서는, 출력 포트의 사용 / 불사용 상황을 나타내는  $N$  bit의 정보만을 전송하면 좋다.

【0058】 5) 계속되고, 크로스바 스위치 101은,  $0 \leq i \leq N-1$ 인 입력  $i$ 와, 입력  $i$ 의 메모리 로케이션  $(h \bmod (N+1))$ 로부터 판독된 어드레스의 출력 포트와의 사이에 크로스바 커넥션을 확립한다.

【0059】 6)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄링 된 송신 입출력 큐의 선두 (HOL) 패킷을 스위치 코어에 확립된 커넥션을 통하여 선택된 출력 라인에 송신한다.

【0060】 짝수개 입력의 파이프라인 RRGs

상술했던 것처럼, 입력 포트 수가 홀수개의 경우, 각 입력은, 해당 입력이 미래의 타임 슬롯을 위한 출력 포트를 선택한 최후의 것이라하면, 갱신된 집합  $O_{k,h}$  인접 입력에 전송하지 않고, 이것에 의해 해당 타임 슬롯의 스케줄링을 완료한다. 그렇지만, 이 홀수 개용의 알고리즘을 짝수의 경우에 직접 적용하면, 몇 개인가의 입력은 여러의 미래의 타임 슬롯에 대하여 스케줄링을 행하고, 역으로, 다른 입력은 완전히(전혀) 스케줄링을 실행하지 않다 되고 버린다. 그 때문에, 입력 포트 수가 짝수개의 스위치를 제어한 점은, 상술한 파이프라인 기법을 수정한다.

【0061】 상술한 정리 1의 증명으로부터, 제어 정보를 블록한 대신에 지연시키는 것으로, 입력이 짝수개의 경우에 대응할 수 있다고 추론할 수 있다. 입력이 짝수개의 경우는, 각 입력은  $N$  개의 타임 슬롯으로 일회만 집합  $O_{h,h}$  차에 보내지 않고, 다음 타임 슬롯에 있어, 해당 입력은 전의 타임 슬롯으로부터 지연된 집합  $O_{h,h}$  송출한다. 입력  $i$ 가 지연된 집합  $O_{h,h}$  전송한 때, 현재의 집합  $O_{k,h}$ 는 전송하지 않는다. 따라서, 해당 입력  $i$ 는  $k$ 번째의 타임 슬롯에 있어 출력을 선택한 최후의 입력으로 된다.

【0062】 정리 2 입력 수  $N$ 이 짝수로,  $(\text{정수} - k) \bmod N$ 의 입력이  $k-2$ 번째의 타임 슬롯에 있어서 집합  $O_{h,h}$  지연시키고,  $k-1$ 번째의 타임 슬롯에 있어 전송한다면,  $(\text{정수} - k) \bmod N$ 의 입력은  $k$ 번째의 타임 슬롯의 스케줄을 완료시킨다.

【0063】 (증명)  $(\text{정수} - k) \bmod N$ 의 입력  $i$ 가  $k-2$ 번째의 타임 슬롯에 있어서 집합  $O_{h,h}$  지연시키고,  $k-1$ 번째의 타임 슬롯에 있어 집합  $O_{k,h}$  대응으로 지연되어진 집합  $O_{h,h}$  전송한 것이라고 가정한다. 이 경우,  $(k-1-j)$ 번째의 타임 슬롯에 있어,  $(i+j-1) \bmod N$ 의 입력은 집합  $O_{m,h}$  지연시키고,  $(i+j) \bmod N$ 의 입력은 지연되어진 집합  $O_{h,h}$  다음 입력으로 전송한다.  $(k-1-j)$ 번째의 타임 슬롯에 있어,  $(i-j) \bmod N$ 의 입력은  $k$ 번째의 타임 슬롯을 예약하고, 집합  $O_{k,h}$  전송한다. 단,  $N$ 이 짝수로,  $0 \leq j \leq N/2 - 1$ 이라고 한 경우,  $i-j \neq i+j \bmod N$  및  $i-j \neq i+j-1 \bmod N$ 이다.

【0064】  $(i - N/2) \bmod N$ 의 입력은,  $k$ 번째의 타임 슬롯을 어느 입력도 예약하지 않도록,  $(k-1 - N/2)$ 번째의 타임 슬롯에 있어 집합  $O_{k,h}$  격납한다.  $(k-2 - N/2)$ 번째의 타임 슬롯에 있어,  $(i - N/2) \bmod N$ 의 입력은,  $k$ 번째의 타임 슬롯을 예약한다.  $N/2 + 2 \leq j \leq N$ 으로서  $(k-1 - N/2)$ 번째의 타임 슬롯에 있어,  $(i-j+1) \bmod N$ 의 입력은  $k$ 번째의 타임 슬롯을 예약하고, 집합  $O_{k,h}$  전송한다. 단,  $N$ 이 짝수로,  $i-j+1 \neq i+j \bmod N$  및  $i-j+1 \neq i+j-1 \bmod N$ 이다.

【0065】 따라서  $k$ 번째의 타임 슬롯의 스케줄링은, 입력  $i$ 에 대응한 유저  $i$ 에 의하고 종료될 때까지 중단한 일 없게 파이프라인에 의하고 진행한다. 이 스케줄이  $(k-N-1)$ 번째의 타임 슬롯으로  $(i+1) \bmod N$ 의 유저 (입력)에 의하고 시작됐다고 말할 것은, 해당 입력이  $(k-N-2)$ 번째의 타임 슬롯에 있어 제어 정보 (즉, 집합  $O$ )의 전송을 지연시켰기 때문이다.

(증명 종)

【0066】그림 3은, 짝수 입력수의  $4 \times 4$  크로스바 스위치를 이용한 경우의 RRGs 스케줄링의 일례를 나타내는 타이밍 차트이다. 그림 3으로는, 4개의 입력  $I_0 \sim I_3$ 과, 그것들 입력이 출력 포트를 선택한 타임 슬롯  $T_1, T_2 \dots$ 라는 관계가 나타나고 있다.

【0067】그림중이 굵은 종선 301은, 그 전의 입력으로 스케줄링이 완료되고 다음 입력으로부터 새로운 스케줄링을 시작한 것을 나타내고, 사선 부 302는 제어 정보O가 지연되고 있는 것을 나타낸다.

【0068】상술했던 것처럼,  $h$ 번째의 타임 슬롯 (예를 들면 현재의 타임 슬롯)에 있어서 RRGs의 동작을 설명한다.  $O_{k,h}$ 는  $k$ 번째의 타임 슬롯의 이용 가능한 출력의 집합을 나타낸다.  $k(i, h) > 0$ 은, 입력  $i$ 가  $h$ 번째의 타임 슬롯에 있어 확보한 출력의 타임 슬롯을 나타내고,  $i_h = (\text{정수} - N - 1 - h) \bmod N$ 은  $h$ 번째의 타임 슬롯으로 새롭게 스케줄링을 시작한 입력을 나타낸다. 또,  $k(i, h) = 0$ 은,  $h$ 번째의 타임 슬롯으로의 입력  $i$ 의 동작이 억제된 것을 의미한다. 스케줄러는 적절한 초기화가 필요하고, 초기화 기간은  $N$  타임 슬롯 계속된다. 초기화가 처음의 타임 슬롯  $T_1$ 로 시작된다고 가정하면, 초기화는  $k(0,1) = k(1,1) = \dots = k(N-1,1) = 0$ , 정수  $= N+2$ 라고 설정한 것에 의하고 시작된다. 즉, 그 이후 갱신되지 않는 한 처음의  $N$  타임 슬롯은 모든 입력의 동작이 억제된다.

【0069】그림 1에 나타내는 시스템에 입각해서 설명하면, 다음과 같이 정식화할 수 있다.

【0070】1)  $O_{h+N+1} = \{0, 1, \dots, N-1\}$ ,  $0 \leq i \leq N-1$ 인  $i$ 가 집합  $\{i_h, (i_h + 1) \bmod N\}$ 의 요소가 아니다면 하여,  $k(i_h, h) = h + N + 1$ ,  $k(\bmod(i_h + 1) \bmod N, h) = k(i_h, h - 2)$ , 및  $k(i, h) = k((i - 1) \bmod N, h - 1)$ .

【0071】2)  $0 \leq i \leq N-1$ 인 입력  $i$ 는, 패킷을 송신해야 할 출력  $j$ 를 출력 집합  $O_{k(i,h)}$ 로부터 RR 방식으로 선택한다. 단,  $k(i, h) \neq 0$ 이다. 또한,  $0 \leq i \leq N-1$  및  $i \neq i_h$ 인 입력  $i$ 는 선행한 타임 슬롯에 있어  $(i - 1) \bmod N$ 의 입력으로부터 집합  $O_{k(i,h)}$ 을 수신하고 있다. 선택된 출력  $j$ 는, 수신한 집합  $O_{k(i,h)}$ 로부터 제외된다.

【0072】3)  $0 \leq i \leq N-1$ 인 입력  $i$ 는, 선택된 출력  $j$ 의 어드레스를 커백션 메모리의 메모리 위치  $k(i, h) \bmod (N + 1)$ 에 기억한다. 대응한 수신 입출력 큐로부터 라인 선두의 HOL 패킷이 별개의 송신 입출력 큐에 옮겨진다.

【0073】4)  $0 \leq i \leq N-1$ 로  $i \neq (i_h - 2) \bmod N$ 인 입력  $i$  (현실 아비타)는, 다음 입력  $(i + 1) \bmod N$  (후속한 아비타)에 집합  $O_{k(i,h)}$ 을 전송한다.  $(i_h - 2) \bmod N$ 의 입력은 집합  $O_{k((i_h - 2) \bmod N, h)}$ 을 1 타임 슬롯 부분만 지연시켰던 후, 전송한다.

【0074】5) 계속되고, 크로스바 스위치 101은,  $0 \leq i \leq N-1$ 인 입력  $i$ 와, 입력  $i$ 의 메모리 위치  $(h \bmod (N + 1))$ 로부터 판독된 어드레스의 출력 포트와의 사이에 크로스바 커백션을 확립한다.

【0075】6)  $0 \leq i \leq N-1$ 인 각 입력  $i$ 에 관하여, 스케줄 된 송신 입출력 큐의 선두 (HOL)의 패킷을 스위치 코어에 확립된 커백션을 통하여 선택된 출력 라인에 송신한다.

【0076】멀티 캐스트 스케줄링

본 발명은 더욱 멀티 캐스트 기능을 갖는다. 멀티 캐스트 패킷은 별개의 큐에 격납되고 도착 순서로 처리된다 (FCFS: first come first served). 각 큐는, 그 HOL 패킷의 수신인을 나타내는 멀티 캐스트 비트맵 (BM)을 갖는다. 가장 간단한 버전으로는, 멀티 캐스트 큐는 유니캐스트 큐보다도 우선되고 처리된다.  $h$ 번째의 타임 슬롯에 있어서 멀티 캐스트의 경우, 다음 동작이 부가된다.

【0077】1)  $0 \leq i \leq N-1$ 인 입력  $i$ 는,  $j \in O_{k(i,h)}$  BM <sub>$i$</sub> 를 충족시키는 모든 출력  $j$ 를 선택한다. HOL 멀티 캐스트 패킷은,  $k$ 번째의 타임 슬롯에 있어, 선택된 모든 출력 포트에 송신된다.

【0078】2)  $O_{k(i,h)}$  BM <sub>$i$</sub> 가 공 집합이라면 유니캐스트 큐가 동작한다. 그 밖의 경우는, 선택된 출력  $i \in O_{k(i,h)}$  및 BM <sub>$i$</sub> 로부터 제외된다.

【0079】3) BM <sub>$i$</sub> 가 하늘이라면, HOL 멀티 캐스트 패킷이 멀티 캐스트 큐로부터 삭제된다.

【0080】스위치 컨트롤러의 실장

도 4는,  $N \times N$  및 크로스바 스위치의 제어를 행한 아비타 부의 구성을 나타내는 블록도이다.  $N$ 개의 입력 포트 ( $0 \sim (N - 1)$ )에 각각 대응하고, 입력 모듈  $IM_0 \sim IM_{N-1}$ , 파이프라인 처리를 실행한 RR 아비타 콘트롤러  $ARRB_0 \sim ARRB_{N-1}$ , 및 커백션 메모리  $M_0 \sim M_{N-1}$ 이 마련되어 있다. RR 아비타 콘트롤러  $ARRB_0 \sim ARRB_{N-1}$ 은, 본 발명에 의한 파이프라인 처리를 행하기 위해 (때문에) 링 위에 접속되고 있다.

【0081】각 입력 모듈  $IM$ 은, 도착한 패킷을 논리적으로 분리한 큐 (수신 입출력 큐)에 격납한다. 상술했던 것처럼, 각 수신 입출력 큐에 의하고, 해당 입력 포트가 어느 특성의 출력 포트에 패킷을 출력하도록 연관되고 있다. 입력 모듈  $IM$ 은, 도착 패킷의 출선 리퀘스트 RQ에 의하고 각 수신 큐에 격납한 도착 패킷의 관리를 행함과 동시에, 출선 리퀘스트 RQ를 연관된 RR 아비타 콘트롤러  $ARRB$ 에 송출한다.

【0082】RR 아비타 콘트롤러  $ARRB$ 는, 입력한 출선 리퀘스트 RQ에 응답하고, 이용 가능 출선 정보 O로부터 미래의 타임 슬롯으로 이용 가능한 출력 포트의 1개를 선택하고, 선택된 출력 포트 (출 철사, 전선의 굵기를 표시하는 번호 호) GR을 연관된 입력 모듈에 되돌린다. 상술한 파이프라인의 초기화 프로세스에 의하고, 어느 타임 슬롯에 있어도, RR 아비타 콘트롤러  $ARRB$ 가 동일한 타임 슬롯에 대하여 중복되고 송신 스케줄링을 행하지 않는 것이 보증된다.

【0083】각 입력 모듈 IM은, 더욱, 미래의 타임 슬롯에 있어 출력 포트의 예약 (스케줄링)에 성공한 패킷을 별개의 송신 입출력 큐로서 격납한다. 또, R R 아비타콘토로라 A R B는, 대응한 커백션 메모리 M의 특정 장소에 스케줄링 결과를 기록한다. 메모리의 어드레스는, 패킷이 스케줄링 된 타임 슬롯에 의하고 결정된다.

【0084】R R 아비타콘토로라 A R B는, 있는 타임 슬롯으로 아직도 예약되고 있지 않는 모든 출력 포트를 나타내는 제어 정보를 후속한 R R 아비타콘토로라 A R B에 통지한다. 보다(부터) 정확하게는, 그 제어 정보에 의하고 예약이 끝난 출력에 대한 출선 리퀘스트가 금지된다. 만약 제어 정보를 다음 포트에 전송하지 않는 것이 있으면, 그 R R 아비타콘토로라에 계속된 다음 R R 아비타콘토로라는 미래의 타임 슬롯으로 임의의 출력을 선택할 수 있는 것이 된다.

【0085】커백션 메모리에 기록된 스케줄에 근거하고, 패킷은 입력 모듈로부터 스위치 코어 101을 통하여 출력 모듈(OM)에 전송된다.

【0086】또한, 본 발명의 변경이나 변형은, 상기 기재 및 교시로부터 당업자에게는 분명할 것 같다. 여기에서는 본 발명의 몇 개인가의 실시 형태만을 기재했지만, 본 발명의 기술적 범위로부터 일탈한 일 없고, 많은 변형예를 한 것은 가능하다.

#### 【0087】성능 비교

먼저, R R G S와 동일한 정도의 복잡함(complexity)을 갖는 다른 프로토콜과의 비교를 행한다. 「복잡함」은, 1개의 스케줄을 완료하는데도 필요한 시간에 의하고 꺾인다. 여기에서는, H O L (Head Of Line), I-TDMA (Interleaved TDMA), SLIP (Iterative round-robin matching with slip), R G S (Random greedy scheduling), 및 R R G S 프로토콜의 비교를 행한다.

【0088】H O L 프로토콜은, 입력 큐를 이용한 스위치를 위한 가장 간단한 프로토콜이다 (M. Karol et al., "Input vs. output queuing on a space-division packet switch," IEEE Transactions on Communications, vol. COM-35, no.12, Dec. 1987, pp. 1347-1356). 각 입력은 적절한 출력에 HOL 패킷의 송신 요구를 송출한다. 요구된 출력은, 라운드 올새 방식으로 입력의 1개에 대해서 송신 허가를 준다. 다음 타임 슬롯으로, 송신 허가를 받았던 입력은 패킷을 대응한 출력에 향하고 송출한다.

【0089】인터리브 TDMA(I-TDMA)에 있어서는, 출력은 입력에 고정화된 방법에 의하고 할당된다 (K. Bogineni et al., "Low-complexity multiple access protocols for wavelength-division multiplexed photonic networks," IEEE Journal on Selected Areas in Communications, vol.11, no.4, May 1993, pp.590-604). 시간이 프레임에 분할되고, 그 프레임의 각 타임 슬롯에 있어 송신 스케줄이 사전에 결정되고 있다. 패킷은, 목적지에 따라 별개의 큐에 격납되고, 그러한 스케줄 된 타임 슬롯으로 송신된다.

【0090】S L I P 프로토콜은, N. McKeown et al. "Scheduling cells in an input-queued switch," Electronic Letters, vol.29, no.25, Dec. 1993, pp.2174-2175에 제안되고 있다. 각 입력은, 송출해야 할 패킷을 갖는 경우, 그것들 패킷의 모든 송출 앞인 출력에 대해서 요구를 낸다. 요구된 출력은, 라운드 올새 방식으로, 요구를 내고 있는 입력의 1개에 대해서 송신 허가를 준다. 여러의 허가를 받았던 입력은, 라운드 올새 방식으로, 허가된 출력의 1개를 선택한다. 라운드 올새 선택은, 전화 선택된 후보자의 차로부터 시작된다.

【0091】R G S 프로토콜은, 라운드 올새 선택을 랜덤 선택으로 대치한 점을 제외하면, R R G S와 유사하다 (D. Guo et al., "Scalable high-speed protocols for WDM optical star networks," IEEE INFOCOM'94). 컨트롤러는 랜덤하게 일련의 입력을 선택하고, 그것들을 비교하지 않았던 출력과 랜덤하게 비교한다. 고속이 되면, R G S는 1 타임 슬롯 안에서 작업을 종료한 것이 할 수 있지 않는다. 그렇지만, 본 발명자들은, 랜덤 선택을 라운드 올새 선택으로 대치한 경우의 영향을 검사하고 그 성능 평가를 행했다.

【0092】그림 5는, R R G S, R G S, H O L, SLIP 및 I-TDMA를 실장한 경우, 제공된 교통 부하에 대한 평균 패킷 지연을 각각 나타내는 그래프이다. 이론 특성과 시뮬레이션 결과란 잘(자주) 일치하고 있다.

【0093】그림 6 (A) 및 (B)은, R R G S, R G S, SLIP 및 I-TDMA를 실장한 경우, 고정 부하 0.8 및 0.9에 있어서 패킷 지연의 상보 분포 함수를 각각 나타내는 그래프이다. 플롯 된 곡선은 시뮬레이션 결과에 근거한다. 대부분(거의)의 부하에 대해서, R R G S는 SLIP 및 I-TDMA보다도 성능이 꽤 뛰어나다.

【0094】그림 7 (A)은, R R G S, R G S, SLIP 및 I-TDMA를 실장하고 교통 부하가 비 단조로운 조건하에서, 입출력 큐 (i, j)의 그룹 G1로의 평균 패킷 지연을 각각 나타내는 그래프이고, 그림 7 (B)은, 동조 건하로 다른 입출력 큐의 그룹 G2로의 평균 패킷 지연을 각각 나타내는 그래프이고, 그림 8 (A)은, 동조 건하로 다른 입출력 큐의 그룹 G3로의 평균 패킷 지연을 각각 나타내는 그래프이고, 그림 8 (B)는, 동조 건하로 다른 입출력 큐의 그룹 G4로의 평균 패킷 지연을 각각 나타내는 그래프이다.

#### 【0095】

【발명의 효과】본 발명은, 입력 버퍼부 고속 스위치를 위한 파이프라인·라운드 올새·스케줄링 방법을 제공한다. 본 발명에 의한 R R G S 프로토콜으로는, 그림 5~그림 8에 나타냈던 것처럼, 동일한 복잡함에 있어, 다른 프로토콜보다도 패킷의 평균 지연 시간이 짧다. 또, 패킷 지연 분포가 크게 넓어지고 있지 않다. 교통 부하가 비 단조로운 경우, 다른 프로토콜과 비교하면, 부하가 적은 큐의 지연 시간은 길어지지만, 부하가 큰 큐의 지연 시간은 훨씬 짧아지고 있다.

#### 【도면의 간단한 설명】

【그림 1】본 발명에 의한 초고속 교환 시스템의 개념적 아키텍처를 나타내는 블록도이다.

【그림 2】출수 입력수의 5×5 크로스바 스위치를 이용한 경우의 R R G S 스케줄링의 일례를 나타내는 타이밍 차트이다.

【그림 3】 짝수 입력수의  $4 \times 4$  크로스바 스위치를 이용한 경우의 RRG S 스케줄링의 일례를 나타내는 타이밍 차트이다.

【그림 4】  $N \times N$  및 크로스바 스위치의 제어를 행한 아비타 부의 구성을 나타내는 블록도이다.

【그림 5】 RRG S, RGS, HOL, SLIP 및 I-TDMA를 실장한 경우, 제공된 교통 부하에 대한 평균 패킷 지연의 시뮬레이션 결과와 이론치를 각각 나타내는 그래프이다.

【그림 6】 (A) 및 (B) 은 , RRG S, RGS, SLIP 및 I-TDMA를 실장한 경우, 고정 부하 0.8 및 0.9에 있어서 패킷 지연의 상보 분포 함수를 각각 나타내는 그래프이다.

【그림 7】 (A) 는 , RRG S, RGS, SLIP 및 I-TDMA를 실장하고 교통 부하가 비 단조로운 조건하에서 , 입출력 큐 (i, j) 의 그룹 G 1 로의 평균 패킷 지연을 각각 나타내는 그래프이고, (B) 는 , 동조 건하로 다른 입출력 큐의 그룹 G 2 로의 평균 패킷 지연을 각각 나타내는 그래프이다.

【그림 8】 (A) 는 , 동조 건하로 다른 입출력 큐의 그룹 G 3 로의 평균 패킷 지연을 각각 나타내는 그래프이고, (B) 는 , 동조 건하로 다른 입출력 큐의 그룹 G 4 로의 평균 패킷 지연을 각각 나타내는 그래프이다.

【부호의 설명】

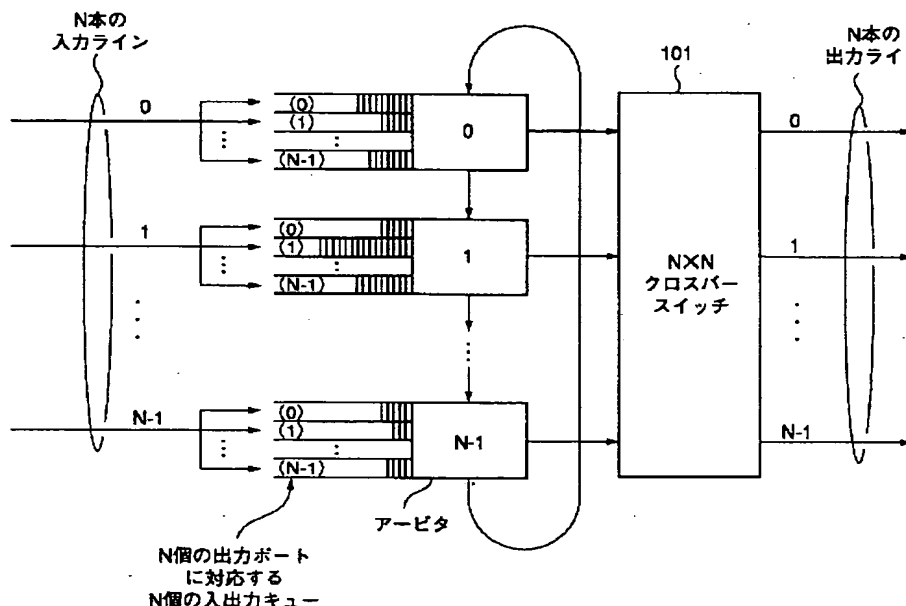
101  $N \times N$  크로스바 스위치

IM 입력 모듈

ARB R R아비타 및 파이프라인·컨트롤러

M 커백션 메모리

【그림 1】

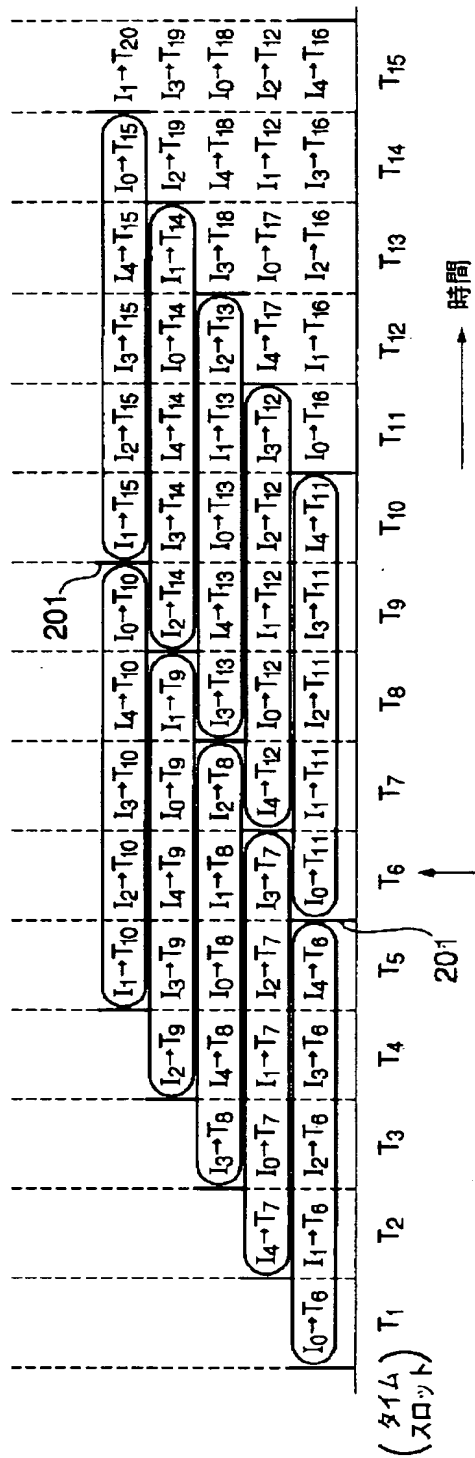


【그림 2】



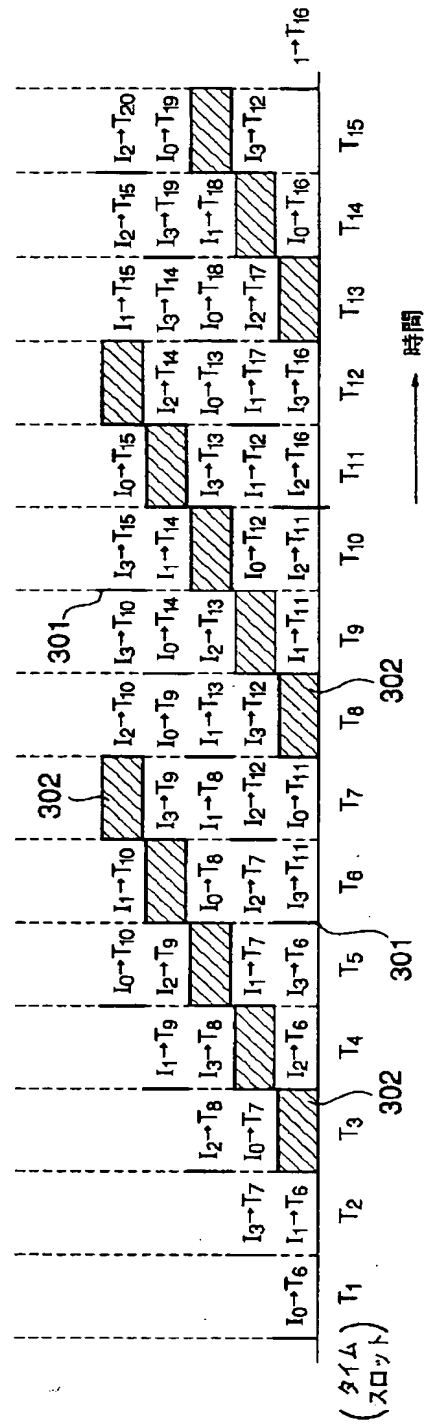
【図 3】

N=5

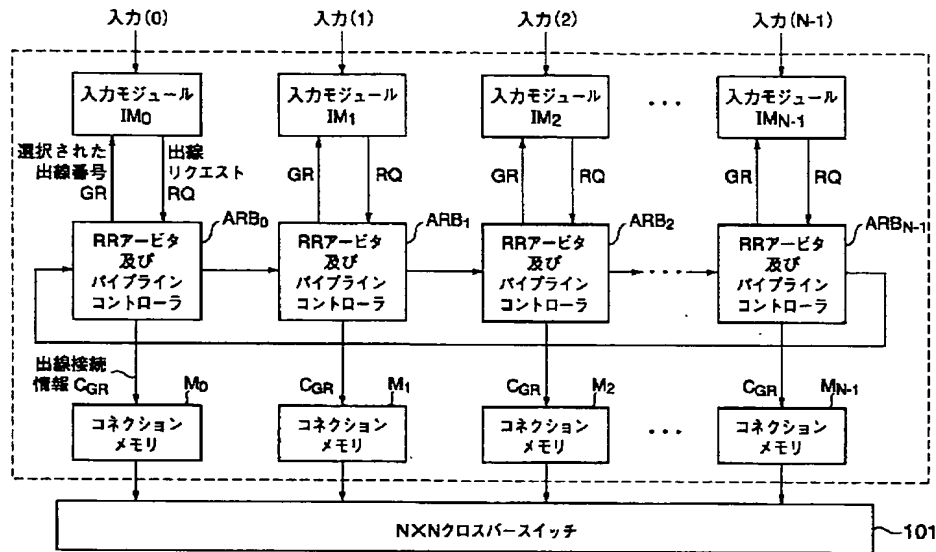


T1~T5で入力I0~I4により  
予約された、タイムスロット

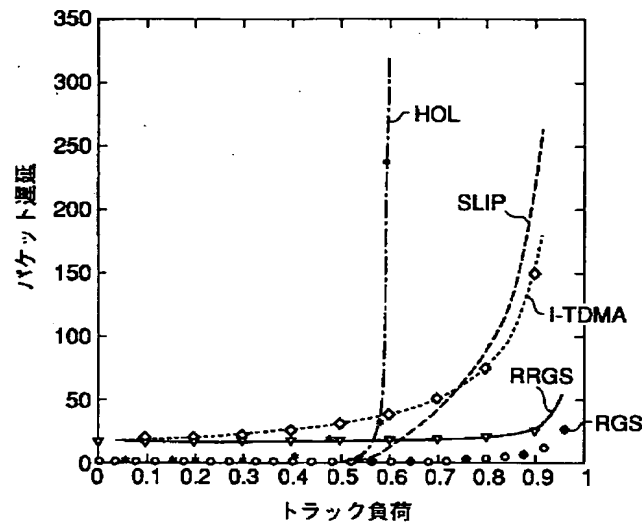
N=4



[그림 4]

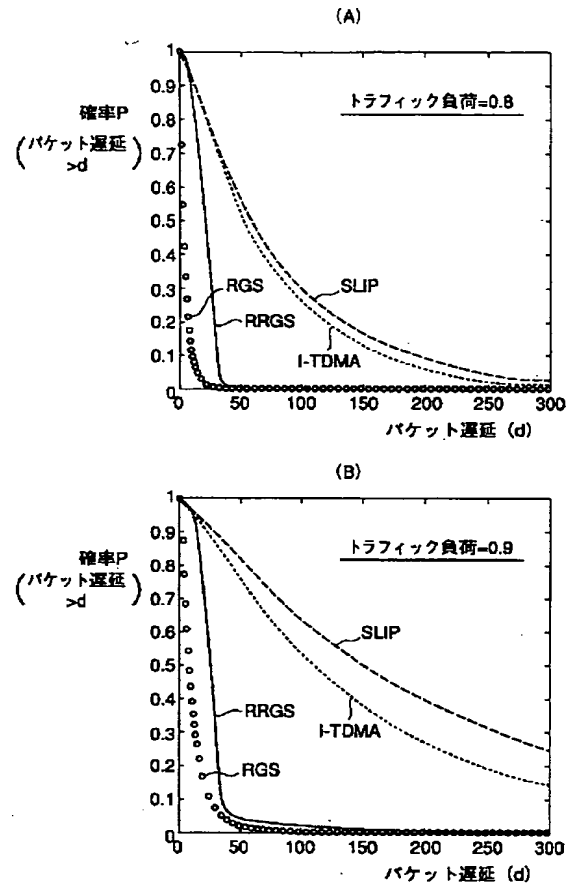


【그림 5】

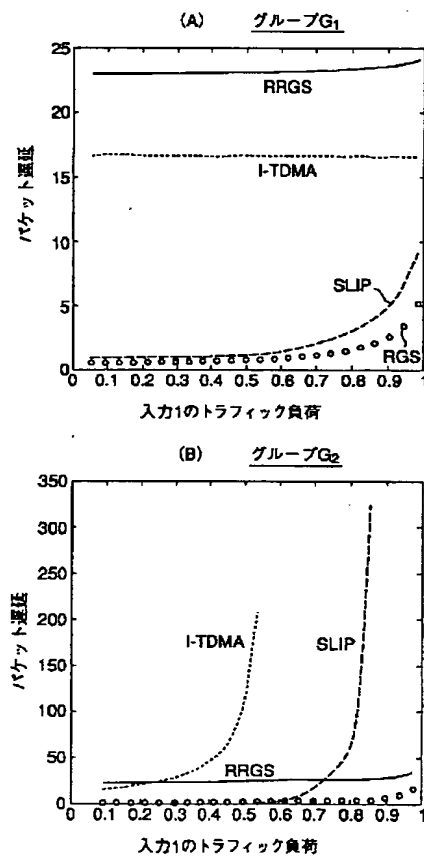


RRGS	{	シュミレーション: —
	{	理論値: ▽
RGS	{	シュミレーション: ○
	{	理論値: +
HOL	{	シュミレーション: —
	{	理論値: *
SLIP	{	シュミレーション: - - -
I-TDMA	{	シュミレーション: ·····
	{	理論値: ◇

【그림 6】



【 그림 7 】



【 그림 8 】